ARMY RESEARCH LABORATORY

A Comparison of Various Types of Head-Related Transfer Functions for 3-D Sound in the Virtual Environment

Douglas S. Savick

ARL-TR-1605 MAY 1998

Fastrak® is a registered trademark of Polhemus, a Kaiser Aerospace and Electronics Company.

Pentium® is a registered trademark of Intel® Corporation.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425

ARL-TR-1605 May 1998

A Comparison of Various Types of Head-Related Transfer Functions for 3-D Sound in the Virtual Environment

Douglas S. Savick
Human Research & Engineering Directorate

Abstract

Simulation using virtual reality (VR) is becoming an effective tool for the Army in training soldiers to do their required tasks. In VR, the human operator can interact with a wide variety of computer-generated worlds developed from real or imaginary scenarios or both. The training that a soldier receives by simulation is usually cost effective to the Army and in a number of cases is safer for the individual than training in the real environment.

Three-dimensional (3-D) sound in the virtual environment (VE) provides a more realistic simulation of acoustic environments compared to diotic (mono) or dichotic (stereo) sound presentation. The major benefit of using 3-D sound is that an individual can determine the sound source direction. When sounds that are perceived to have direction and sights that represent virtual objects that produce the sounds are provided through a head-mounted display, a person can monitor and identify sources of information from all possible locations.

The purpose of this study was to determine if 3-D sound generated by a Tucker-Davis Technologies' 3-D sound system could enhance the "realism" or fidelity of the VE. The main objective of the study was to determine if an individual could distinguish the direction of a sound source within a reasonable degree of accuracy. Three-dimensional sound is produced by using a mathematical representation of the filtering characteristics of the pinnae provided through head-related transfer functions (HRTFs). The HRTFs can be developed by recording a generated broadband sound using a probe microphone in the ear canal and subsequently dividing the Fourier transform of the recorded sound by that of the generated sound. When digital filtering techniques are used, HRTFs can be applied to sounds through headphones. When an arbitrary sound is filtered with HRTF-based filters, the sound should appear to come from specified virtual locations outside the earphones. Ideally, every person should have his or her own unique or "matched" HRTFs to generate localized sound. Because the development of matched HRTFs is time consuming, generic or "unmatched" HRTFs are used to satisfy a broad range of listeners. This study featured a comparison between using matched HRTFs versus generic HRTFs.

The results indicate that the average localization errors for the baseline scenario and the scenario that used the generic HRTFs were small and close in value. The difference, although statistically significant, has therefore no practical importance. The average localization error for custom HRTFs, however, was approximately 2.5 times larger than that of the baseline scenario. These results were contrary to what should be expected.

ACKNOWLEDGMENTS

I would like to thank Dr. Kim Abouchacra, Dr. Tomasz Letowski, and Mr. Tim Mermagen for their guidance, technical support, and expertise in auditory research. I would also like to thank the subjects for their enthusiastic participation.

CONTENTS

A COMPARISON OF VARIOUS TYPES OF HEAD-RELATED TRANSFER FUNCTIONS FOR 3-D SOUND IN THE VIRTUAL ENVIRONMENT

INTRODUCTION

Simulation using virtual reality (VR) is becoming an effective tool for the Army in training soldiers to do their required tasks. In VR, the human operator is connected to a computer that can simulate a wide variety of worlds, both real and imaginary. The training that a soldier receives by simulation is usually cost effective to the Army and, in a number of cases, is safer for the individual than training in the real environment. Because funding and safety are major issues in today's Army, interests in VR continue to grow.

Continuous enhancement of existing virtual environments is also of great interest to both the developers and the users. In general, as the virtual environment (VE) becomes more realistic, training becomes more effective. That is, training in high resolution, high fidelity VEs should better prepare soldiers for real-world conditions. Failure to make such enhancements may result in training that does not allow the soldiers to use their full capabilities or lead the users to believe that they have more capabilities than they truly have in a real environment.

The advancements in audio displays have presented an attractive means of enhancing the VE. In any setting, sound is a major component of soldiers' perception of their environment. The localization of sounds guides the listener, making that person more aware of his or her surroundings and increasing his or her sense of security. When quality three-dimensional (3-D) sound is incorporated into the VE, the individual's sense of immersion is greatly enhanced.

BACKGROUND

Key steps to developing a "life-like" simulation have been initiated through the "I-Port" program, which is a collaborative research and development program under the direction of the Human Research and Engineering Directorate (HRED) of the U.S. Army Research Laboratory (ARL). The I-Port program provides a reconfigurable simulator for training in the VE. The Uniport, a prototype simulator for the I-Port program, is a unicycle-type motion platform that allows the individual to move ("pedal") through the VE. It extracts the appropriate amount of energy from the individual as a function of the type of virtual terrain he or she is traveling. The Uniport's successor, the individual soldier mobility simulator (ISMS), provides the individual with a more natural means of walking by supporting him or her with a robotic footpad under each foot. The footpads passively follow the movement of each foot to provide support directly

related to the virtual terrain on which the person is walking. The addition of a helmet-mounted display (HMD) or a rear projection system, commonly known as a walk-in simulated environment (WISE), allows the soldier to see the terrain and environment with which he or she will interact. Finally, real-time simulation of the environment is provided through the Naval Postgraduate School Network (NPSNET). The physical and visual feedback from these simulators provide a means for immersing an individual in a synthetic environment to some level of fidelity.

Complete immersion of an individual in the VE is difficult to accomplish. Currently, no VR system can fully facilitate for every motion that a human makes or simulate every stimulus that a human can perceive. However, in the past few years, major advancements have been made in simulating signals that affect the human senses. One such advancement is the use of 3-D sound in the VE. Incorporating 3-D sound in the VE provides a more realistic simulation of acoustic environments compared to a diotic (mono) or dichotic (stereo) sound presentation. The major benefit of using 3-D sound is that an individual can determine the sound source direction and distance in space. Individuals hear sounds in their natural environment from every direction in 3-D space, and the untrained ear can determine the sound source's direction within a 10° localization error (Wenzel, 1992). By generating sounds from corresponding virtual objects throughout the VE, a person has the potential to monitor and identify sources of information from all possible locations.

Producing localized sound, however, is a complex process. Sounds produced in the real world are perceived from a certain location based on intensity and time differences between the ears and the spectral shaping by the individual's pinnae (outer ear) (Wenzel, 1992). Specifically, the outer ear acts as a direction-dependent sound filter to aid the listener in "pinpointing" sound sources. A mathematical representation of the filtering characteristics of the pinnae is provided by the head-related transfer function (HRTF): "the complex valued free field transfer function from a sound source in a certain direction to the eardrum." (Bronkehorst, 1995). The HRTF can be developed by recording a generated broadband sound at a particular spatial location, through a probe microphone in the ear canal and subsequently dividing the Fourier transform of the recorded sound by that of the generated sound (Wightman & Kistler, 1989). When digital filtering techniques are used, the HRTFs can be applied to sounds before those sounds are presented through headphones. When an arbitrary sound is filtered with HRTF-based filters, the sound will be perceived as originating from specified virtual locations outside the earphones.

As with all human attributes, a wide degree of variability exists in the anatomical characteristics of the human head and ear. Specifically, the pinnae vary in size, shape, and even location on the human head. Even minute differences in these features can influence a listener's perception of sound source location. For the most accurate spatial representation of sound in the VE, every individual should use a unique or "matched" set of HRTFs to represent his or her perception of sound. Developing HRTFs for every person is possible but is time consuming. For this reason, many commercial applications use generic or "unmatched" HRTFs that are sufficient for most listeners to perceive 3-D acoustic images. For military applications, however, precision is important and generic HRTFs may not provide enough resolution. Tucker-Davis Technologies, Inc. (TDT) has recently developed a procedure to match an individual with a set of HRTFs that closely represents his or her spatial perceptions. During this procedure, the individual's HRTFs are customized by having the person select from a large library of previously recorded HRTFs that he or she chose as the best representative for localizing sound at specified locations.

OBJECTIVE

The objectives of the study were to determine (1) if an individual could localize a sound produced by TDT's 3-D sound system with reasonable accuracy for use in a VE, and (2) whether individualized HRTFs are necessary. The study featured a comparison between using matched HRTFs versus generic HRTFs. Localization accuracy in the real environment was used as a baseline.

METHODS

Equipment

Instrumentation used in this study was comprised of a psycho-acoustic signal-processing system, two computers, and standard input-output equipment. Both custom-made and commercial software was used to run the experiments. The specific makes and models of the equipment are as follow:

a. Tucker-Davis Technologies System II processing system (see Figure 1)
b. International Business Machine (IBM) personal computer (PC)-Pentium®
c. Pioneer speakers
d. AKG (not an acronym) acoustic earphones (Model 240 DF)
e. Polhemus Fastrak® tracking system

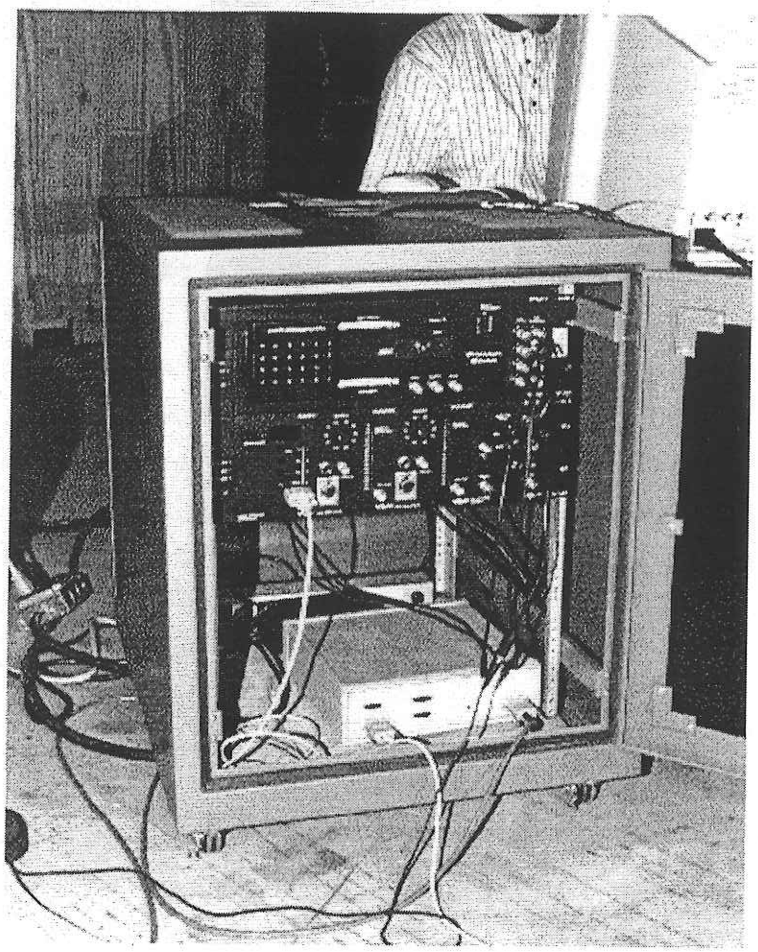

Figure 1. Tucker-Davis Technologies System II processing system and Polhemus Fastrak®.

The sound stimulus consisted of a single shot from an M16. The stimulus was a good representative of broadband sounds that are easily localized by listeners with normal hearing in a natural environment. The level of the stimulus was adjusted to a comfortable and consistent level at the subject's ear. The distance to the sound source in both the real and virtual environments was 2.5 meters.

Customizing Software

Software developed by TDT provided the user with a means for developing his or her own customized HRTFs. Each individual listened through earphones to a generated broadband sound played at various known locations around the perimeter of his or her head. For each location, the listeners selected the "best" HRTF representation. When they had completed the procedure, their selected HRTFs were used as part of the study.

The software begins by displaying on the monitor the location where the sound will be played. For example, the first location is 90° right of due forward (at the right ear). The user will select any number on the keypad, and a previously recorded HRTF assigned to that number will display the sound at that location. The listener, however, will notice that not all of the selections will appear to be coming from exactly 90° right of due forward. As mentioned before, differences in a person's features can influence his or her perception of the sound source location. When one listens to 3-D sound through HRTFs recorded from other persons' ears, he or she is in a sense listening through someone else's ears. The listeners will select the HRTF that they believe to be the best representation of the sound coming from that specific location. Once the selection has been made, the process is repeated at the next location, 120° right of due forward, for a total of 12 locations at 30° increments. The computer software then interpolates the HRTFs for the locations in between the 30° increments.

Testing Environment

Testing took place in the ARL hostile environment simulator (HES), which is a 57-foot by 44-foot by 22-foot chamber. A subject was seated on a chair (a) centered within an array of speakers (5 meters in diameter) that were located every 30° at ear level (see Figure 2) or (b) with earphones placed over the ears (see Figure 3). Background noise levels in the chamber were kept below recommended noise levels for audiometric testing using earphone simulation (American National Standards Institute, 1991).

Subjects

Twenty subjects (10 males and 10 females) between the ages of 18 and 40 were exposed to three different scenarios in which they tried to localize sounds from any direction at ear level (0° to 360° azimuth). Each subject had hearing better than or equal to 20 dB hearing level (HL) at audiometric frequencies from 250 through 8000 Hz (American National Standards Institute, 1996), otoscopically normal ears, and no history of otologic pathology.

Procedure

The sound stimulus was displayed at 12 locations (30° increments) for each of the three scenarios. The locations of the sound were randomly presented to the listener for each scenario. In each scenario, the subject was seated in a revolving chair, wearing either a headband or earphones supporting a Polhemus Fastrak® tracking sensor (static accuracy of 0.15° root mean square). The

subject was blindfolded when appropriate to prevent visual objects (such as speakers in the room) from playing a role in his or her response to the stimulus.

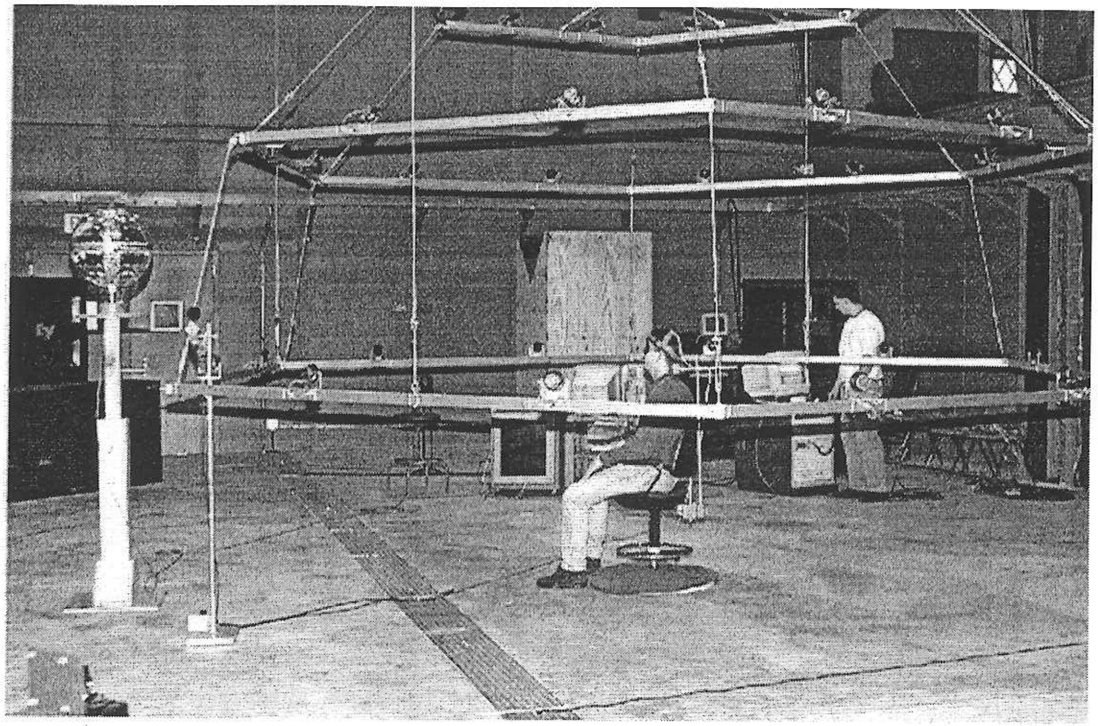

Figure 2. Speaker array, 360° azimuth, for baseline localization accuracy.

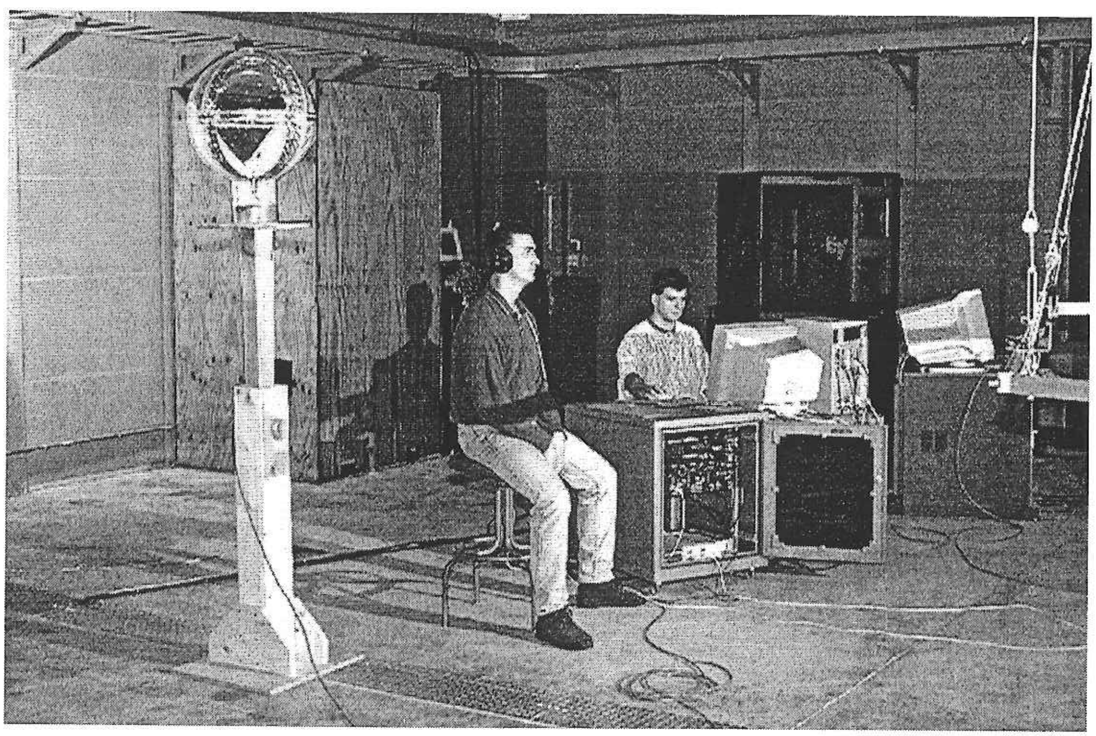

Figure 3. Three-dimensional sound display through headphones using various HRTFs.

One scenario served as a baseline. The subject, in this scenario, sat in the chair and listened for sounds produced through the 360° azimuth speaker array. He or she was blindfolded and asked to determine the direction of the produced sound by turning and facing the perceived sound. The sound was repeated every 3 seconds from the selected location until the individual acknowledged verbally to stop the sound because he or she was facing the direction from which he or she perceived the sound to be coming. Once the sound was stopped, the direction of the subject was recorded from the output of the tracking sensor.

The other two scenarios followed the same testing procedure as in the baseline scenario except the sound stimulus was produced through earphones from the 3-D sound system. The 3-D sound system produced the sound stimulus using generic HRTFs for the one scenario and customized HRTFs for the other scenario. Each scenario was fully completed before another scenario began. A sample of the presentation order for one subject for all three scenarios is shown in Table 1.

Table 1

Sample Configuration for a Subject to Localize a Sound
Played at the Various Locations for Each Scenario
(Twelve locations, every 30°, were randomly selected.)

Location number	1	2	3	4	5	6	7	8	9	10	11	12
Speaker array (degrees)	0	150	90	180	210	330	240	120	60	300	30	270
Generic HRTFs (degrees)	330	210	300	270	60	120	150	240	90	180	0	30
Custom HRTFs (degrees)	270	30	60	240	0	210	300	120	330	180	90	150

RESULTS AND DATA ANALYSIS

The recorded data consisted of each subject's localization error (in degrees) at each location where the sound was displayed for the three scenarios. The data are shown in Appendix A. A multivariate repeated measures analysis of variance (MANOVA) was conducted on localization accuracy. The two main factors were scenario and gender. The analysis of the data collected from 10 males and 10 females did not reveal any statistically significant differences

between female and male for all scenarios ($p = 0.143$). Therefore, all the data were combined and the conclusion derived from this analysis applied to both female and male listeners. The MANOVA also showed that there was a statistical difference between the localization errors of each scenario ($p < 0.002$). Post hoc contrast analysis indicated that all three data sets significantly differed one from another ($p < 0.05$)

Table 2 provides a basic comparison between the results produced by the TDT system (generic and custom HRTFs) and the real life sound (baseline). Table 2 shows the average localization error of all 20 subjects for each of the three scenarios determined by averaging the subjects' average localization errors for each of the three scenarios. In other words, each subject's localization error was averaged over the 12 locations measured for each scenario. Then the averaged localization errors for each subject were averaged together to determine the total average localization error for all subjects in each scenario. The standard deviations were determined in the same manner.

Table 2

Average Localization Error and Standard Deviation for Each Scenario

	Baseline	Generic	Custom
Average (degrees)	3.37	4.89	8.21
Standard deviation	0.97	1.68	2.83

The results in Table 2 show very little difference between the baseline scenario and the scenario that used the generic HRTFs. The difference, although statistically significant, has therefore no practical importance. The average localization error for custom HRTFs, however, is approximately 2.5 times larger than the baseline. These results are contrary to what should be expected. The average localization error for the custom scenario should have been much closer to the baseline average localization error, provided that the customizing software was able to generate a compatible match of each subject's HRTFs. This indicates that the customization procedure developed by TDT was too difficult to use by some of the subjects.

The average localization error for the custom scenario does appear, though, to still be less than an individual's 10° localization error noted by Wenzel. However, when we observe the

individual's average localization error for each scenario, as seen in Figure 4, we see a large variation for localization error using the custom HRTFs (s = 2.83) with 5 of 20 subjects having a localization error of more than 10°. In addition, it was noted that a number of individuals stated that it was more difficult to determine the exact location of the sound source during this portion of the study.

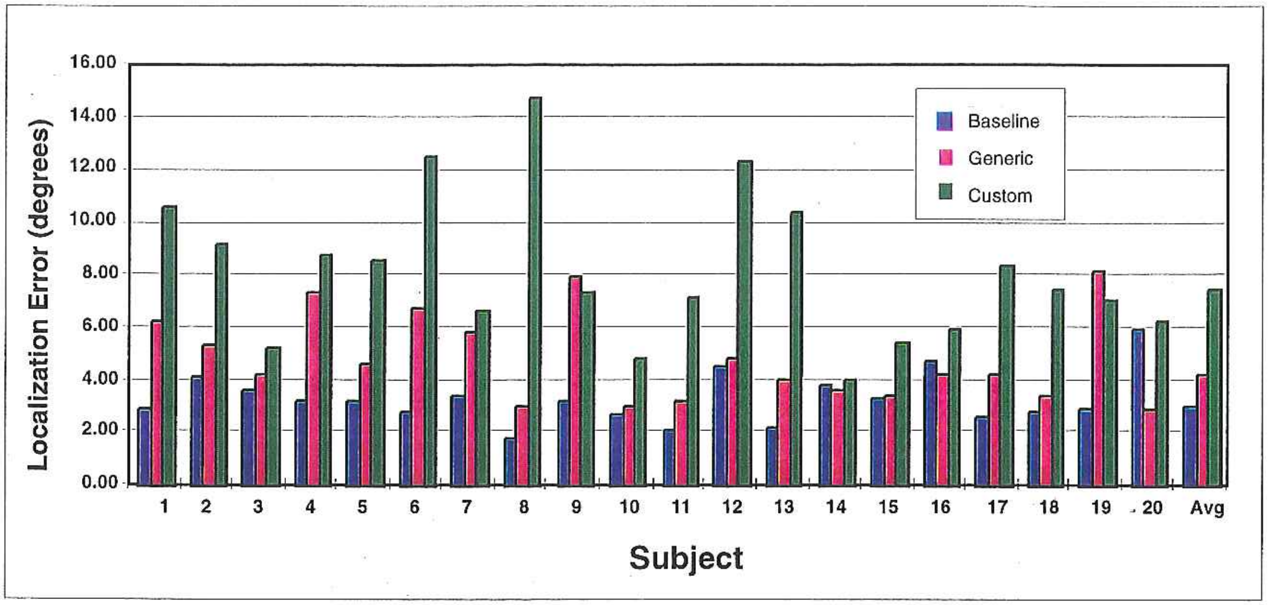

Figure 4. Average localization error for individuals for each scenario.

In addition, it was also of interest to look at the average localization error at each location for each scenario, represented in Figure 5. If certain locations showed a significant error in comparison to the rest of the locations (the data were not close to being linear), there would be a need for further investigation of whether individuals have a greater localization error at particular locations from the initial direction that they were facing. Figure 5, however, shows no discrepancies that would lead one to consider that the precision of matching was direction dependent.

CONCLUSIONS

A VR training simulator must be versatile and easy to use. Effective simulator training cannot always allow the individual the necessary time to be fitted with some type of customized HRTFs, be it semi-custom, as was done in this study, or a fully measured HRTF that requires

hours of recorded sounds through microphones placed on a person's ears. Generic (one size fits all) HRTFs are the most attractive approach if it provides sufficient accuracy for localization.

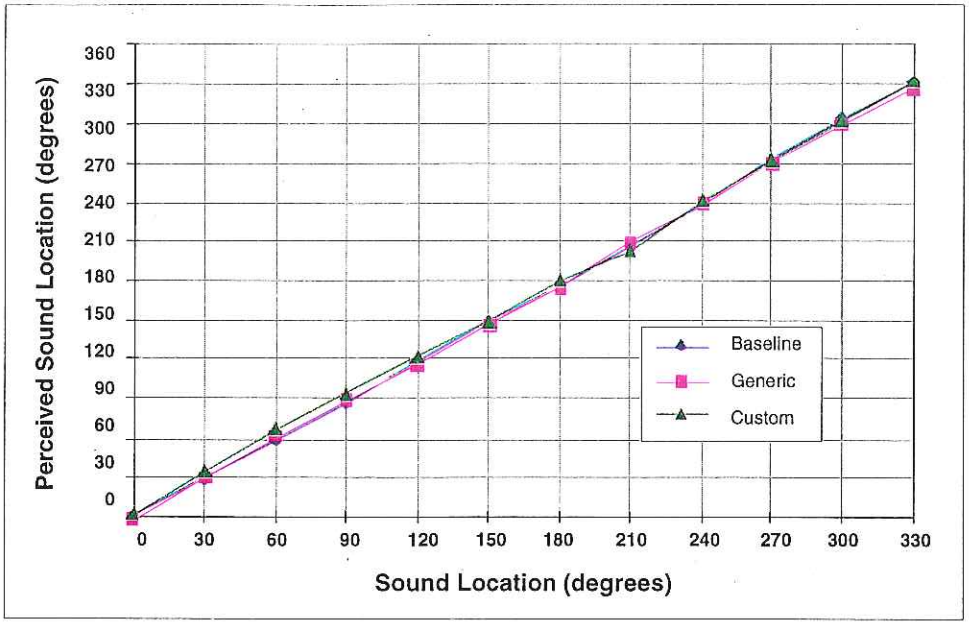

Figure 5. Average perceived sound at each location in each scenario.

The results of this study show that the 3-D sound produced through both the generic and customized HRTFs provided localized sound with an average localization error of less than 10°. However, for this particular study, the generic HRTFs appear to be the better choice for producing 3-D sound. One reason is that in comparison to the baseline, the average localization error recorded using generic HRTFs differed by less than 2°. Another reason is that there was no need to determine an individual's HRTFs before exposing him or her to various 3-D sound cues.

It is important to point out that the semi-custom approach is an attractive means for developing accurate HRTFs for individuals. Currently, though, semi-customizing does require anywhere from 15 minutes to 2 hours to develop a set of HRTFs, depending on the thoroughness of the individual. Being more thorough, however, did not necessarily mean better accuracy for this study. If the software for semi-customizing is improved and a broader range of previously measured HRTFs can be used, an individual may be able to closely replicate HRTFs similar to those he or she would obtain if the HRTFs were measured in a fraction of the time.

REFERENCES

American National Standards Institute (1991). Maximum permissible ambient noise levels for audiometric test rooms (ANSI S3.1-1991). New York: Author.

American National Standards Institute (1996). Specifications for audiometers (ANSI S3.6-1996). New York: Author.

Bronkehorst, A.W. (1995). Localization of real and virtual sound sources. Journal of the Acoustical Society of America, 98(5), 2542-2553.

Wenzel, E.M. (1992). Localization in virtual acoustic displays. Presence: Teleoperators and Virtual Environments, 1, 80-107.

Wightman, F.L., & Kistler, D.J. (1989). Headphone simulation of free-field listening: I. stimulus synthesis. Journal of the Acoustical Society of America, 85, 858-867.

APPENDIX A

LOCALIZATION ERROR DATA

LOCALIZATION ERROR DATA

Females
Baseline Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	3	0	1	1	0	1	0	1	1	4	1.20	1.32
30°	3	1	4	0	3	2	6	2	1	2	2.40	1.71
60°	3	6	5	6	1	2	6	2	5	0	3.60	2.27
90°	3	8	5	3	6	1	0	3	0	4	3.30	2.58
120°	4	6	2	1	5	2	2	3	1	4	3.00	1.70
150°	1	8	4	2	4	1	2	0	1	2	2.50	2.32
180°	7	8	5	3	11	3	9	1	4	2	5.30	3.30
210°	0	5	7	3	2	5	11	5	10	3	5.10	3.45
240°	1	3	3	3	2	2	1	2	7	0	2.40	1.90
270°	3	1	1	8	0	5	1	1	4	3	2.70	2.45
300°	3	2	7	7	3	5	2	2	3	4	3.80	1.93
330°	5	2	1	3	3	5	2	1	3	5	3.00	1.56

Females
Generic HRTFs Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	2	3	1	2	0	2	3	2	6	1	2.20	1.62
30°	6	2	3	13	3	19	6	1	10	4	6.70	5.70
60°	10	4	10	7	13	7	2	4	5	0	6.20	3.99
90°	6	0	6	6	3	9	4	1	3	6	4.40	2.72
120°	6	7	4	4	13	15	2	7	9	2	6.90	4.38
150°	0	1	1	8	9	4	8	2	10	5	4.80	3.74
180°	8	2	5	12	1	3	10	0	12	8	6.10	4.51
210°	24	17	3	12	0	0	18	1	5	1	8.10	8.90
240°	4	4	2	15	0	8	6	6	0	2	4.70	4.47
270°	4	17	5	0	5	1	1	2	19	1	5.50	6.84
300°	3	8	9	8	4	8	6	7	8	2	6.30	2.45
330°	2	0	3	2	5	6	5	4	9	5	4.10	2.51

Females
Custom HRTFs Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	7	5	6	0	21	2	1	3	6	1	5.20	6.07
30°	15	27	13	21	24	15	2	22	2	2	14.30	9.52
60°	18	3	8	2	4	15	8	22	2	13	9.50	7.15
90°	16	5	2	4	9	8	7	15	8	2	7.60	4.84
120°	20	15	2	16	3	22	1	10	1	4	9.40	8.25
150°	14	6	8	16	1	8	6	3	16	4	8.20	5.39
180°	7	2	13	3	9	5	8	4	8	12	7.10	3.67
210°	12	26	0	13	14	21	16	32	5	3	14.20	10.13
240°	2	16	6	0	14	6	5	16	17	2	8.40	6.64
270°	1	3	1	7	1	25	14	30	12	10	10.40	10.24
300°	1	0	5	18	1	13	13	8	5	4	6.80	6.07
330°	15	3	0	6	2	11	0	13	7	1	5.80	5.55

Males
Baseline Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	0	8	1	4	3	0	2	5	3	1	2.70	2.50
30°	2	4	2	5	3	4	0	3	4	4	3.10	1.45
60°	1	0	2	5	2	0	6	1	6	10	3.30	3.30
90°	1	1	2	0	1	13	0	0	3	5	2.60	3.98
120°	0	4	0	5	1	5	0	4	6	9	3.40	3.06
150°	1	5	1	4	11	3	1	2	3	9	4.00	3.46
180°	1	1	2	0	2	0	0	3	3	7	1.90	2.13
210°	7	5	2	0	1	3	5	1	1	8	3.30	2.79
240°	3	4	1	5	4	3	4	2	1	6	3.30	1.64
270°	1	6	3	4	9	7	3	3	3	5	4.40	2.37
300°	6	7	8	9	3	11	5	4	1	4	5.80	3.01
330°	3	10	3	5	1	8	6	6	2	4	4.80	2.78

Males
Generic HRTFs Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	3	7	1	5	2	3	1	5	15	1	4.30	4.27
30°	2	2	0	3	4	0	2	2	6	0	2.10	1.91
60°	2	10	7	2	9	2	5	1	12	1	5.10	4.12
90°	2	6	3	3	5	7	6	2	15	4	5.30	3.83
120°	4	4	7	5	4	3	7	0	7	3	4.40	2.22
150°	6	0	3	6	1	10	4	1	5	2	3.80	3.05
180°	5	4	0	7	2	9	7	9	9	5	5.70	3.09
210°	1	0	0	3	2	6	0	0	11	5	2.80	3.61
240°	6	8	5	6	3	5	10	10	2	2	5.70	2.95
270°	2	3	2	3	3	1	5	6	0	0	2.50	1.96
300°	3	7	5	1	5	6	4	2	8	3	4.40	2.22
330°	4	8	13	0	2	0	0	4	9	10	5.00	4.71

Males
Custom HRTFs Differences

	Subject											
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Avg.	SD
0°	9	22	5	0	2	5	2	13	6	2	6.60	6.64
30°	7	23	11	4	5	6	3	10	5	4	7.80	5.94
60°	5	25	32	3	9	2	6	2	7	8	9.90	10.22
90°	10	0	18	9	8	1	5	14	15	2	8.20	6.21
120°	12	7	17	2	12	8	12	4	8	19	10.10	5.36
150°	13	9	6	8	0	14	6	19	21	8	10.40	6.38
180°	10	12	17	11	13	2	12	7	1	16	10.10	5.34
210°	4	9	11	8	0	22	21	2	11	7	9.50	7.29
240°	2	19	7	3	12	4	8	5	0	0	6.00	5.89
270°	11	7	1	3	2	6	15	13	4	2	6.40	4.99
300°	4	29	11	3	7	9	20	5	10	5	10.30	8.21
330°	14	11	7	2	5	4	7	4	7	13	7.40	4.03

NO. OF COPIES	ORGANIZATION
2	ADMINISTRATOR DEFENSE TECHNICAL INFO CENTER ATTN DTIC DDA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1	DIRECTOR US ARMY RESEARCH LABORATORY ATTN AMSRL CS AL TA REC MGMT 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LABORATORY ATTN AMSRL CI LL TECH LIB 2800 POWDER MILL RD ADELPHI MD 207830-1197
1	DIRECTOR US ARMY RESEARCH LABORATORY ATTN AMSRL CS AL TP TECH PUB BR 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR ARMY AUDIOLOGY & SPEECH CENTER WALTER REED ARMY MED CENTER WASHINGTON DC 20307-5001
1	WALTER REED ARMY INST OF RSCH ATTN SGRD UWI C (COL REDMOND) WASHINGTON DC 20307-5100
1	COMMANDER US ARMY RESEARCH INSTITUTE ATTN PERI ZT (DR E M JOHNSON) 5001 EISENHOWER AVENUE ALEXANDRIA VA 22333-5600
1	USA BIOMEDICAL R&D LABORATORY ATTN LIBRARY FORT DETRICK BUILDING 568 FREDERICK MD 21702-5010
1	HQ USAMRDC ATTN SGRD PLC FORT DETRICK MD 21701
1	COMMANDER USA AEROMEDICAL RESEARCH LAB ATTN LIBRARY FORT RUCKER AL 36362-5292
1	CHIEF ARMY RESEARCH INSTITUTE AVIATION R&D ACTIVITY ATTN PERI IR FORT RUCKER AL 36362-5354
1	US ARMY NATICK RD&E CENTER ATTN STRNC YBA NATICK MA 01760-5020
1	US ARMY TROOP SUPPORT CMD NATICK RD&E CENTER ATTN BEHAVIORAL SCI DIV SSD NATICK MA 01760-5020
1	US ARMY TROOP SUPPORT CMD NATICK RD&E CENTER ATTN TECH LIBRARY (STRNC MIL) NATICK MA 01760-5040
1	STRICOM 12350 RESEARCH PARKWAY ORLANDO FL 32826-3276
1	GOVT PUBLICATIONS LIBRARY 409 WILSON M UNIVERSITY OF MINNESOTA MINNEAPOLIS MN 55455
1	HUMAN FACTORS ENG PROGRAM DEPT OF BIOMEDICAL ENGINEERING COLLEGE OF ENGINEERING & COMPUTER SCIENCE WRIGHT STATE UNIVERSITY DAYTON OH 45435
3	DARPA L STOTTS J PENNELLA B KASPAR 3701 N FAIRFAX DR ARLINGTON VA 22203-1714
1	ARL HRED AVNC FIELD ELEMENT ATTN AMSRL HR MJ (R ARMSTRONG) PO BOX 620716 BLDG 514 FT RUCKER AL 36362-0716
1	ARL HRED MICOM FIELD ELEMENT ATTN AMSRL HR MO (T COOK) BUILDING 5400 ROOM C242 REDSTONE ARSENAL AL 35898-7290

NO. OF COPIES	ORGANIZATION
1	ARL HRED USAADASCH FLD ELEMENT ATTN AMSRL HR ME (K REYNOLDS) ATTN ATSA CD 5800 CARTER ROAD FORT BLISS TX 79916-3802
1	ARL HRED ARDEC FIELD ELEMENT ATTN AMSRL HR MG (R SPINE) BUILDING 333 PICATINNY ARSENAL NJ 07806-5000
1	ARL HRED ARMC FIELD ELEMENT ATTN AMSRL HR MH (J JOHNSON) BLDG 1109B 3RD FLOOR FT KNOX KY 40121-5215
1	ARL HRED CECOM FIELD ELEMENT ATTN AMSRL HR ML (J MARTIN) MYER CENTER RM 3C214 FT MONMOUTH NJ 07703-5630
1	ARL HRED FT BELVOIR FIELD ELEMENT ATTN AMSRL HR MK (P SCHOOL) 10115 GRIDLEY ROAD SUITE 114 FORT BELVOIR VA 22060-5846
1	ARL HRED FT HOOD FIELD ELEMENT ATTN AMSRL HR MV (E SMOOTZ) HQ TEXCOM BLDG 91012 RM 111 FT HOOD TX 76544-5065
2	ARL HRED NATICK FIELD ELEMENT ATTN AMSRL HR MQ (M FLETCHER) ATTN SSCNC A (D SEARS) USASSCOM NRDEC BLDG 3 RM R-140 NATICK MA 01760-5015
1	ARL HRED FT HUACHUCA FLD ELEMENT ATTN AMSRL HR MY (B KNAPP) GREELY HALL (BLDG 61801 RM 2631) FORT HUACHUCA AZ 85613-5000
1	ARL HRED FT LEAVENWORTH FLD ELE ATTN AMSRL HR MP (D UNGVARSKY) TPIO ABCS 415 SHERMAN AVE RM 327 FT LEAVENWORTH KS 66027-1344
1	ARL HRED FLW FIELD ELEMENT ATTN AMSRL HR MZ (A DAVISON)* 320 ENGINEER LOOP STE 166 FT LEONARD WOOD MO 65473-8929

NO. OF COPIES	ORGANIZATION
1	ARL HRED OPTEC FIELD ELEMENT ATTN AMSRL HR MR (D HEADLEY) PARK CENTER IV RM 1450 4501 FORD AVENUE ALEXANDRIA VA 22302-1458
1	ARL HRED SC&FG FIELD ELEMENT ATTN AMSRL HR MS (L BUCKALEW) SIGNAL TOWERS RM 207 FORT GORDON GA 30905-5233
1	ARL HRED STRICOM FIELD ELEMENT ATTN AMSRL HR MT (A GALBAVY) 12350 RESEARCH PARKWAY ORLANDO FL 32826-3276
1	ARL HRED TACOM FIELD ELEMENT ATTN AMSRL HR MU (M SINGAPORE) BLDG 200A 2ND FLOOR WARREN MI 48397-5000
1	ARL HRED USAFAS FIELD ELEMENT ATTN AMSRL HR MF (L PIERCE) BLDG 3040 RM 220 FORT SILL OK 73503-5600
1	ARL HRED USAIC FIELD ELEMENT ATTN AMSRL HR MW (E REDDEN) BLDG 4 ROOM 332 FT BENNING GA 31905-5400
1	ARL HRED USASOC FIELD ELEMENT ATTN AMSRL HR MN (F MALKIN) HQ USASOC BLDG E2929 FORT BRAGG NC 28307-5000
1	US ARMY RSCH DEV STDZN GP-UK ATTN DR MICHAEL H STRUB PSC 802 BOX 15 FPO AE 09499-1500
	<u>ABERDEEN PROVING GROUND</u>
2	DIRECTOR US ARMY RESEARCH LABORATORY ATTN AMSRL CI LP (TECH LIB) BLDG 305 APG AA
1	LIBRARY ARL BLDG 459 APG-AA

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1998	Final

4. TITLE AND SUBTITLE

A Comparison of Various Types of Head-Related Transfer Functions for 3-D Sound in the Virtual Environment

5. FUNDING NUMBERS

AMS Code 622716.H700011
PR: 1L162716AH70
PE: 6.27.16

6. AUTHOR(S)

Savick, D.S. (ARL)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Human Research & Engineering Directorate
Aberdeen Proving Ground, MD 21005-5425

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARL-TR-1605

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Simulation using virtual reality (VR) is becoming an effective tool for the Army in training soldiers to do their required tasks. In VR, the human operator can interact with a wide variety of computer-generated worlds developed from real or imaginary scenarios or both. The training that a soldier receives by simulation is usually cost effective to the Army and in a number of cases is safer for the individual than training in the real environment.

Three-dimensional (3-D) sound in the virtual environment (VE) provides a more realistic simulation of acoustic environments compared to diotic (mono) or dichotic (stereo) sound presentation. The major benefit of using 3-D sound is that an individual can determine the sound source direction. When sounds that are perceived to have direction and sights that represent virtual objects that produce the sounds are provided through a head-mounted display, a person can monitor and identify sources of information from all possible locations.

The purpose of this study was to determine if 3-D sound generated by a Tucker-Davis Technologies' 3-D sound system could enhance the "realism" or fidelity of the VE. The main objective of the study was to determine if an individual could distinguish the direction of a sound source within a reasonable degree of accuracy. Three-dimensional sound is produced by using a mathematical representation of the filtering characteristics of the pinnae provided through head-related transfer functions (HRTFs). The HRTFs can be developed by recording a generated broadband sound using a probe microphone in the ear canal and subsequently dividing the Fourier transform of the recorded sound by that of the generated sound. When digital filtering techniques are used, HRTFs can be applied to sounds through headphones. When an arbitrary sound is filtered with HRTF-based filters, the sound should appear to come from specified virtual locations outside the earphones. Ideally, every person should have his or her own unique or "matched" HRTFs to generate localized sound. Because the development of matched HRTFs is time consuming, generic or "unmatched" HRTFs are used to satisfy a broad range of listeners. This study featured a comparison between using matched HRTFs versus generic HRTFs. (continued on reverse)

14. SUBJECT TERMS

3-D sound, localization, HRTF, virtual environment

15. NUMBER OF PAGES

25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Item 13 (continued)

The results indicate that the average localization errors for the baseline scenario and the scenario that used the generic HRTFs were small and close in value. The difference, although statistically significant, has therefore no practical importance. The average localization error for custom HRTFs, however, was approximately 2.5 times larger than that of the baseline scenario. These results were contrary to what should be expected.